Development of Sediment Quality Objectives for California Bays and Estuaries

Workplan for:
Development of Benthic Community Condition Indicators

October 19, 2004

Background

The California State Water Resources Control Board's effort to develop sediment quality objectives focuses on maintaining healthy natural living resources in the sediments in enclosed bays and estuaries. A central aspect of this effort is the development of benthic indicators that will directly assess the condition of a major component of the living resources in bays and estuaries. Such indicators will enable the legislative intent to be addressed directly. They will also help avoid the ambiguity and necessarily uncertain assumptions about causal mechanisms that typically accompany the use of more indirect chemical and toxicity measures alone. The project's overall strategy of using benthic indicators in combination with sediment chemistry and toxicity testing approaches will result in a robust assessment of sediment condition. This workplan addresses technical issues involved in developing the benthic indicator(s) themselves, while other workplans in this package address the other approaches (sediment chemistry, toxicity) that will make up the other portions of the overall sediment quality objectives. The remainder of this section reviews the rationale for focusing on benthic indicators and summarizes their current use in regulatory and assessment regimes. Subsequent sections provide detail on the specific technical tasks proposed.

Benthic communities are found almost universally in aquatic soft sediments and are indicators of choice for monitoring and assessing anthropogenic effects for two main reasons. First, they possess many attributes considered desirable in indicator organisms, including:

- they have limited mobility, making them indicative of impacts at the site where they are collected
- they include several different animal phyla and classes and are therefore sensitive to many types of impacts and respond to different impacts in different ways
- their life-histories are short enough that the effects of one-time impacts disappear within a year but long enough to integrate the effects of multiple impacts occurring within seasonal time scales
- living within the bottom sediments, they are readily exposed to sediment contamination, high sediment organic carbon resulting from eutrophication, and low bottom dissolved oxygen, the three most common anthropogenic impacts in bays and estuaries.

Second, they are important components of aquatic food webs, transferring carbon and nutrients from suspended particulates in the water column to the sediments by filter feeding and serving as forage for bottom-feeding fishes.

Despite these appealing characteristics, benthic infaunal monitoring data are maximally useful in a regulatory context only when they can be interpreted in relation to scientifically valid criteria or thresholds that distinguish "healthy" from "unhealthy" benthic communities. While reducing complex biological data to index values has disadvantages, the resulting indices remove much of the subjectivity associated with data interpretation. Such indices also provide a simple means of communicating complex information to managers, tracking trends over time, and correlating benthic responses with stressor data (Dauer *et al.* 2000, Hale *et al.* 2004).

During the past decade, several scientifically valid measures of marine and estuarine benthic community condition, often called benthic indices, have been developed for regulatory use and provide a useful starting point for the current effort. Most of these are applicable on a regional basis and identify regional reference conditions and deviations attributable to anthropogenic disturbances. These benthic indices generally fall into three categories: (1) discriminant analysis approaches used in Gulf (Engle *et al.* 1994, Engle and Summers 1999) and northeast Atlantic (Paul *et al.* 2001) coast estuaries; (2) Index of Biotic Integrity (IBI) approaches used in Chesapeake (Weisberg *et al.* 1997, Alden *et al.* 2002) and San Francisco (Thompson and Lowe 2004) Bays and the southeast (Van Dolah *et al.* 1999), mid-Atlantic (Llanso *et al.* 2002), and several California (Fairey *et al.* 1996, Anderson *et al.* 1997, 1998, Fairey *et al.* 1998, Jacobi *et al.* 1998, Anderson *et al.* 2001, Hunt *et al.* 2001) estuaries; (3) a Benthic Response Index (BRI) approach used on the southern California mainland shelf (Smith *et al.* 2001) and extended later to coastal bays (Smith *et al.* 2003).

Such applications have indicated that changes in benthic communities occur at chemical concentrations an order of magnitude lower than those which affect sediment toxicity tests (Long *et al.* 2001, Hyland *et al.* 1999, 2003). Partly as a result, benthic indices are increasingly accepted by regulators and incorporated into regulatory processes. The U.S. Environmental Protection Agency's guidance for biocriteria development (Gibson *et al.* 2000) recognizes all three types of benthic indices and the agency included benthic assessments in a recent report on nationwide coastal condition to Congress (U.S. Environmental Protection Agency 2004). In Maryland and Virginia, the Chesapeake B-IBI is one of the measures used to report on the condition of Chesapeake Bay waters under sections 305(b) and 303(d) of the Clean Water Act. In California, benthic indices were one of the factors used by the State Water Control Board to designate toxic hot-spots (California State Water Resources Control Board 1999) and by the San Diego Regional Water Quality Control Board to make clean-up decisions for three toxic hot spots in San Diego Bay (Exponent 2002, Southern California Coastal Water Research Project and Space and Naval Warfare Systems Center San Diego 2004). Due to the presence of benthic communities in good condition as measured by the BRI and other reasons, Santa Monica Bay, which previously was listed as impaired under section 303(d) of the Clean Water Act due to sediment concentrations of six metals, was removed from the list in 2003. The BRI has also been used in southern California to assess the extent of bottom area supporting unhealthy benthic communities since 1994 (Bergen *et al.* 1998, Bergen *et al.* 2000, Ranasinghe *et al.* 2003).

Although this history of theoretical development and practical, regulatory application provides a solid basis for using benthic indices to develop sediment quality objectives, at present there are several impediments to applying them statewide in California's bays and estuaries. First, the number of unique habitats and benthic assemblages that exist and the corresponding number of benthic indices to be developed is unknown; species and abundances of benthic organisms vary naturally from habitat to habitat and comparisons to determine altered states should vary accordingly. Second, different benthic indices have been used in California at different times and different places and results cannot be compared across regions because the various indices have not yet been rigorously compared and intercalibrated; statewide sediment quality objectives should be equally protective irrespective of region. Third, initial development of each existing benthic index was constrained by data limitations and they would all benefit from refinement with additional data as well as independent validation. In addition, there is a lack of knowledge of the effects of differences in 1) sampling procedures traditional in different regions, 2) habitat factors such as seasonality and sediment type, and 3) accuracy of identification of benthic organisms on performance of California benthic indices.

The work plan that follows describes the objectives, approach, and details of tasks intended to overcome these impediments and develop benthic community condition indicators for statewide sediment quality objectives for California bays and estuaries.

Objectives

The scientific and technical objectives of this work plan are:

- Define habitat strata for development of benthic indicators. Benthic indicators are developed, refined and validated separately for each habitat stratum because species and abundances of benthic organisms vary naturally from habitat to habitat and comparisons to determine altered states should vary accordingly.
- Develop one or more indicators of benthic community response for all habitat strata defined in the first objective for which sufficient data are available. The indicators may include traditional benthic community measures, refined versions of existing benthic indices, and other community or population measures.
- Identify assessment thresholds corresponding to ecological thresholds of impacts on aquatic life for each indicator of benthic community response developed in the second objective.
- Create a consistent set of approaches for California as a whole by comparing benthic indicator results in each habitat stratum for which multiple indicators are available and evaluating their responses at different ecological thresholds of impacts on aquatic life.
- Evaluate responses of the benthic indicators to habitat factors such as seasonality, depth, salinity and sediment type.
- Evaluate the uncertainty in benthic indicator responses due to factors such as
    - Differences in sampling protocols traditionally used in different parts of California,
    - Accuracy of benthic identifications, and
    - Habitat factors such as seasonality, depth, salinity and sediment type.

General Approach

Two types of activities are necessary to provide the technical foundation for the benthic portion of the SQOs: 1) Develop indicators to summarize benthic infaunal data into quantitative and interpretable values, and 2) Describe effective methods for sampling and analysis. This workplan includes three tasks related to indicator development:

> **Task 1: Identify naturally occurring assemblages.** The objective of this task is to define habitat strata for development of benthic indicators by identifying the naturally occurring assemblages in California and the habitat factors that structure them. The species and abundances of benthic organisms vary naturally from habitat to habitat and, therefore, benthic indicators and definitions of reference condition should vary accordingly.
>
> **Task 2: Refine existing benthic indicators.** The objective of this task is to refine and improve the three benthic indices that are available in California and explore the efficacy of traditional benthic community measures. Recent data will be used to develop and validate a Relative Benthic Index (RBI), an IBI, and a BRI for the enclosed bays and estuaries of southern California and for San Francisco Bay. Traditional benthic community measures commonly used for assessment purposes will also be calculated in these two regions.
>
> **Task 3: Compare and evaluate benthic tools.** The objective of this task is to develop an index application strategy based on ecologically appropriate threshold values. In addition to the three indices, traditional benthic community measures commonly used for assessment purposes will be evaluated in an effort to develop preliminary objectives for areas in which insufficient data are available for full index development. The effects of habitat factors such as seasonality, depth, salinity and sediment type on available benthic indicators and the magnitude and sources of uncertainty benthic indicators will be evaluated as part of this task.

Two tasks are included in the this workplan for development of methods guidance:

> **Task 4: Evaluate field sampling methods.** The sampling gear typically used in California varies regionally and inhibits the comparison of data among studies. The objective of this task is to describe the effects of differences in gear size and sieve size on benthic assessment and benthic community measures.
>
> **Task 5: Develop taxonomy QA procedures.** The reliability of benthic assessment measures depends heavily on accurate sorting, species identifications, and counts. The objective of this task is to develop procedures to document and assure the quality of the taxonomy results.

Work Description

Task 1: Identify Naturally Occurring Assemblages. The species and abundances of benthic organisms vary naturally from habitat to habitat, and comparisons to determine altered states should vary accordingly. Once the habitat factors responsible for assemblage differences are identified, the bays and estuaries will be stratified accordingly and benthic assessment tools developed for each stratum in subsequent tasks.

A process similar to Bergen *et al.* (2001) will used to identify the benthic assemblages that occur naturally in California and the U.S. west coast and the habitat factors that structure them. After eliminating potentially contaminated sites, assemblages will be identified using hierarchical cluster analysis and habitat variables will be tested across dendrogram splits to assess whether the assemblages occupy different habitats.

1.1. Data. Data from studies over broad geographic areas will be used for this task. The U.S. EPA's EMAP study, which uses consistent and compatible methods throughout California and the West Coast, will provide the foundation. Where the EMAP sampling density is insufficient to capture the expected range of natural habitat variation, the EMAP data will be supplemented by regional data collected using compatible methods. Data used for this task will meet three criteria:

- Benthic species abundance data from samples sieved through 1-mm screens.
- Data about the sampling sites will include habitat (depth and sediment grain size distribution), sediment chemistry and amphipod toxicity data in addition to location and benthic species abundance data.
- All data types will have been subjected to quality assurance and quality control measures equivalent to those used by the U.S. EPA's Environmental Monitoring and Assessment Program.

1.2. Data Analysis. Because the objective is to define natural groupings of samples with similar species composition, screening criteria similar to those of Bergen *et al.* (2001) will be used to eliminate potentially contaminated sites from the analysis. For example, samples might be considered potentially contaminated if the mean ERM quotient is more than 0.1 (Long and MacDonald 1998).

Cluster analysis will be conducted using flexible sorting of Bray-Curtis dissimilarity values with β=-0.25 (Bray and Curtis 1957, Lance and Williams 1967, Clifford and Stephenson 1975). Abundances will be cube-root transformed and then standardized by the species mean of values higher than zero to reduce the influence of dominant species. Prior to cluster analysis, species contributing little information will be excluded using numbers of occurrences as a species screening criterion.

The number of habitat-defined assemblages will be determined by sequentially examining each split of the cluster analysis dendrogram, starting at the top, to assess whether each split reflects habitat differentiation. Habitat differentiation will be defined as a significant (Mann-Whitney U-test) difference in habitat variables between the sets of samples defined by the dendrogram split and segregation of more than, for example, 90% of the samples in the split by the significant habitat variables. The habitat variables to be tested will include salinity, depth, fine sediment content, total organic carbon, latitude and longitude.

1.3. Results. The products of this task will include a map showing habitat divisions of the enclosed bays and estuaries. The map will be based on habitat variable differences across dendrogram splits in a hierarchical cluster analysis. A draft manuscript summarizing the results will also be prepared for journal publication.

Task 2: Refine Existing Benthic Tools. Three benthic indices have been developed for bays and estuaries in different parts of California (Table 1). The Benthic Response Index (BRI; Smith *et al.* 2003) was developed for bays and harbors in southern California, the Index of Biotic Integrity (IBI; Thompson and Lowe 2004) was developed for San Francisco Bay, and several versions of the Relative Benthic Index (RBI; Anderson *et al.* 2001) were developed for different southern California bays and harbors. Although sediment chemistry and toxicity data are used during development of most indices, assessments are based only on biological factors such as species abundances and numbers of species.

All three indices are considered preliminary because of data limitations that constrained their initial development. These limitations included: (1) lack of independent data for validation of the index; (2) insufficient data from highly disturbed sites to define the entire range of the impact gradient; and (3) uncertainty in the effect of environmental variables that can affect assemblage distributions regardless of pollution impacts.

Additional data are now available to surpass these limitations in index development efforts for two of the most populous regions of California, viz., southern California and San Francisco Bay. The recent data include highly contaminated sites to better define the polluted end of the impact gradient and, sufficient data to withhold a portion for independent validation of the indices, at least in southern California.

The objective of this task is to refine all three indices and explore the usefulness of benthic community measures traditionally used to assess benthic condition.

2.1. Data. Regional databases will be compiled for southern California and San Francisco Bay, which are the two regions of California that have sufficient data for development of benthic indicators. The databases will be compiled from existing data sets from multiple studies (Table 2) describing benthic species abundances, sediment contaminant concentrations, and sediment toxicity to amphipods for stations in each of the two regions. Southern California data will include benthic data collected with 0.1 m^2 Van Veen grabs and sieved through 1.0 mm screens while San Francisco Bay data will include benthic data only from samples sieved through 0.5 mm screens.

Table 1. Comparison of Benthic Response Index (BRI), Index of Biotic Integrity (IBI) and Relative Benthic Index (RBI) approaches.			
	BRI	**IBI**	**RBI**
General Approach	The BRI is the abundance-weighted average pollution tolerance score of species at a site. Pollution tolerance scores are calculated during index development and applied during assessment.	Reference ranges are identified for indicator variables from a set of reference stations. The IBI score is the number of indicator variables with values outside reference range.	Calculates the ratio of indicator variables to maximum values in the development data. Ratios are combined and scaled to calculate the index.
Biological measures	Species abundances	Number of taxa, Number of amphipod taxa, number of molluscan taxa, and abundances of one or two pollution indicative species.	Number of taxa, number of molluscan and crustacean taxa, abundance of crustacea, and abundances of two negative and three positive indicator species.
Rationale	A species tolerance to pollutant effects can be characterized by the position of its abundance peak on the pollution gradient.	Some characteristics of infaunal communities and certain species abundances change in response to pollution and other disturbances. Each may be used as independent evidence that conditions differ from reference.	Some characteristics of infaunal communities and certain species abundances change in response to pollution and other disturbances. A site may be ranked relative to other sites based on integration of the indicator variables.
Use of pollutant data in index development	Sediment contaminant concentrations (mean ERM quotient) and amphipod mortality in sediment toxicity tests in index development data are used to establish a pollution gradient and derive species pollution tolerance scores.	Sediment contaminant concentrations (mean ERM quotient) and toxicity used to 1) screen indicator variables and 2) select reference sites.	None.
Selection of biotic measures	None. Pollution tolerance scores developed for all species with sufficient data.	Based on reports of sensitivity to disturbance in literature and screening by multiple regressions to assure response to contamination.	Based on reports of sensitivity to disturbance in literature and best professional judgment.
Assessment threshold	Based on loss of species as index values increase.	Two or more indicators outside reference range.	Identified using best professional judgment.

The San Francisco Bay data are all different from those used in Task 1 because of the sieve size difference while the southern California database will include additional data. The southern California data for Task 1 will be supplemented by several localized studies that include samples from potentially polluted areas.

The data used for this task will meet two criteria:

- Data about the sampling sites will include habitat (depth and sediment grain size distribution), sediment chemistry and amphipod toxicity data in addition to location and benthic species abundance data.

- All data types will have been subjected to quality assurance and quality control measures equivalent to those used by the U.S. EPA's Environmental Monitoring and Assessment Program.

Table 2. Index development data (Tentative).

Region	Source	Period	Location	Numbers of sites
Southern California	Bight'98	Summer 1998	Southern California	123
	EMAP	Summer 1999	Southern California	24
	Chollas and Paleta Creek	Summer 2001 to Summer 2002; Quarterly	San Diego Bay	36
	Switzer Creek, Broadway and B Street Piers, Downtown Anchorage	February and July 2003	San Diego Bay	33
	NASSCO and Southwest Marine	Summer 2001 and Summer 2002	San Diego Bay	70
	Santa Ana RWQB	August 2001 and Summer 2003	Anaheim Bay and Huntington Harbor	59
	Bight'03	Summer 2003	Southern California	120
	Targeted BRI development sites	Summer 2003	Southern California	20
			Total	**485**
San Francisco Bay	BADA	1994-1997	San Francisco Bay	9
	BPTCP-92	May 1992	San Francisco Bay	4
	BPTCP-94	Sep 1994	San Francisco Bay	3
	BPTCP-97	Apr or Dec 1997	San Francisco Bay	21
	CISNET	2000-2001	San Francisco Bay	4
	DWR	Jan 1994 – Dec 2001	San Francisco Bay	15
	RMP	1994-2000	San Francisco Bay	9
	RMP-W	Feb-Mar 1995	San Francisco Bay	4
	EMAP-ML	Jul-Aug 2000	San Francisco Bay	17
	EMAP-NO	Jul-Aug 2000	San Francisco Bay	33
			Total	**119**

<u>2.2. Data Analysis.</u> For each of the two regions for which sufficient data exist, BRI, IBI and RBI indices will be developed using the techniques and approaches specific to each index for a random subset of about 2/3 of the data compiled in Subtask 2.1. The other 1/3 of the data will be withheld for one of the index validation studies in Task 3. Existing indices will be refined and, if an index was not previously available, it will be developed. Each index and its assessment thresholds will be developed using independent analyses of the same data by the teams of investigators who developed the original indices. Community measures commonly used for assessment purposes and potential indicator species abundances will also be calculated for potential use as indicators in areas with insufficient data for benthic index development.

<u>2.3. Results.</u> The products of this task will include refined versions of the BRI, IBI and RBI for each of the two regions. It will also include, for every site for which BRI, IBI and RBI values are available, calculations of benthic community measures commonly used for assessment purposes and potential indicator species abundances.

Task 3: Compare and Evaluate Benthic Tools. The overarching objective of this task is to develop a strategy for application of benthic community indicators to California bays and estuaries, based on four data analysis elements:

- Validation of the three benthic indices
- Comparison of the three benthic indices and traditional benthic community indicators,
- Evaluation of the effects of seasonal and habitat factors on performance of the three benthic indices and benthic community indicators, and
- Measuring the magnitude and sources of uncertainty in the benthic indicators.

The analysis elements are described in more detail in Section 3.2.

In addition to developing a benthic index application strategy, this task will also attempt to identify interim benthic indicators for use in areas where indices and data for index development are currently unavailable. The selection of these interim indicators will be partly based on the relationships among benthic indices, community measures commonly used for assessment purposes, and potential indicator species abundances.

<u>3.1. Data.</u> The database developed for Subtask 2.1 will be used for these analyses, with benthic index, benthic community measure, and indicator species abundances computed during Task 2.

<u>3.2. Data Analysis.</u> Results from the four data analysis elements of this task will be used to develop a strategy for application of benthic community indicators to California Bays and Estuaries. The data analysis elements are:

- Indicator Validation. Multiple approaches will be used to validate the three benthic indices and other indicators. We will attempt to use every available means of validation, potentially including evaluating:
    - Performance of the indicators by comparing values with the sediment chemistry and sediment toxicity data for the random 1/3 of data withheld from index development;
    - Behaviors of the indicator values along spatial and temporal pollution gradients;
    - Indicator stability by examining values at repeatedly sampled sites;
    - Correspondence of indicator values with assessments of benthic condition based on species abundance, habitat, and other data by benthic ecologists not involved in index development; and
    - Correspondence of index thresholds with ecologically important thresholds, such as appearances and disappearances of higher taxonomic groups of organisms.
- Indicator Comparison. Correlations among values of the different benthic indicators will be compared for two purposes. First, high correlations among the different indicators, which are based on different premises and function at differing taxonomic and ecological levels, are a form of validation. Second, if benthic community measures traditionally used to assess benthic condition are highly correlated with the benthic indices or sediment contaminant concentrations or toxicity, they might be useful for use in assessments where insufficient data are available for benthic index development. If they are consistently and highly correlated with benthic indices, contaminant concentrations, or sediment toxicity, it may be possible to use them instead of the complicated benthic indices.

- Evaluation of the effects of seasonal, habitat, and other factors on benthic indicator performance. Analyses will be conducted to describe how robust each index is in respect to factors such as (1) seasonality, (2) grain size distribution, and (3) depth.
- Evaluation of the magnitude of variability and uncertainty in the benthic indicators. Analyses will be conducted to estimate the amount of variability in the index values that is associated with variations in sampling, taxonomy (level and accuracy), and calculation of the index.

3.3. Develop Application Strategy. Based on the data analysis results and input from the SWRCB regarding intended policy applications, an application strategy will be developed for each index. SWRCB guidance regarding the desired level(s) of protection to be afforded the benthos, and the relative importance of avoiding false positive or false negative results will be integrated with the data analysis results to establish several numeric thresholds for each index. Ideally, these thresholds will correspond to ecologically relevant changes in community composition shown by the data analyses conducted in Subtask 3.2.

Task 4: Evaluate Field Sampling Methods. Two elements of the sampling protocols used in California to collect benthic organisms vary from study to study. First, the sampler type and area sampled varies among regions and monitoring programs. In southern California and San Francisco Bay, infaunal samples are typically collected using $0.1m^2$ and $0.05m^2$ Van Veen grabs, respectively. Monitoring programs in other areas of the state often use 10 cm diameter diver cores, with a surface area of $0.00785\ m^2$. Larger samplers collect more organisms and more species, but require more time for sorting, identification, and enumeration in the laboratory and, therefore, cost more to process.

Second, the mesh size used to sieve the samples varies. Most monitoring efforts in central and northern California use 0.5 mm-mesh screens, while most southern California studies use 1 mm–mesh screens. A 0.5 mm screen retains more organisms, resulting in increased laboratory processing costs. In many cases, the organisms passing through the larger screen are juveniles of species that are retained in it, but the larger screen can also miss smaller species entirely, particularly if population density is low.

Data analyses will be conducted to determine the nature and magnitude of effects that variations in gear type and sieve size have on benthic community assessment results. Evaluations of gear area effects will be based on data from sites sampled simultaneously with multiple gears while sieve size effects will be based on samples sieved through both 0.5 and 1.0 mm sieves. Effects on the benthic indices as well as benthic community measures commonly used to assess benthic community condition will be evaluated.

This task will develop recommendations about the need for gear standardization in future benthic collections and determine whether benthic indices should vary with collection method, especially when applied to previously collected data. The recommendations will be based on evaluations of the sensitivity of the benthic indices to gear area and sieve size as well as cost-effectiveness.

4.1. Data. This study will leverage five recent or ongoing benthic data collection efforts by adding sampling or sample processing elements that allow the gear and sieve size questions to be addressed:

- Sixty-four samples, collected as part of EPA's EMAP program, will be processed for a second sieve size. These samples were collected in 1999 and 2000 using a $0.1m^2$ Van Veen grab from a stratified random set of locations throughout coastal California. These samples were sieved using both a 0.5 and a 1 mm screen, but only the 1 mm fraction was analyzed by EMAP. This project will analyze the 0.5 mm fraction.
- EMAP is planning to collect samples with a $0.1m^2$ Van Veen grab and a 1 mm screen from 50 randomly selected sites throughout California in August 2004. This sampling effort will be augmented by using a sampler constructed from paired 0.05 m^2 Van Veen grabs, providing a combined surface area of 0.1 m^2. From one side of the paired grab, a 10 cm diameter core (0.00785 m^2) will be used to subsample the 0.05 m^2 area, with the core and the remainder sieved separately. The other side of the grab will be sieved intact. Sediment from all three subsamples will be sieved through 1.0mm and 0.5mm nested sieves resulting in six infaunal subsample fractions from each station. Data from this effort will be combined to yield synoptically sampled areas of 0.00785, 0.05, and $0.1m^2$, each with a 0.5 and 1.0mm sieve fractions, for assessing the effect of sampling protocols.
- The California Central Coast Regional Water Quality Control Board is collecting samples from 30 random locations from the central coast of California during August of 2004 using a $0.1m^2$ Van Veen grab and a 1 mm screen. This sampling effort will be augmented with the same gear and subsampling approach outlined above for EMAP.
- The San Diego Regional Water Quality Control Board is planning to collect benthic samples at nine targeted locations during August 2004 for assessing site-specific effects in San Diego Bay. This sampling effort will be augmented with the same gear and subsampling approach outlined above for EMAP. The San Diego Bay sampling allows us to ensure that a number of stressed benthic community sites are included in the analysis, as such sites are not likely to be well represented in the randomly selected sites from the programs above. Data from stressed sites are critical for the sediment quality objectives development process.
- In San Francisco Bay, 0.5 and 1.0 mm sieve data are available for 103 samples collected with 0.05 m^2 Van Veen grabs. Data from 0.1 m^2 Van Veen and 0.00785 m^2core samplers are also available for 24 EMAP samples collected in summer 1999. No additional processing of these samples is necessary.

Sediment chemistry data will be collected for the collaborative efforts being augmented for this task and are available for some of the existing San Francisco data. The time taken to process each sample in the field and in the laboratory will be recorded to facilitate analysis of cost-effectiveness.

4.2. Data Analysis. Several measures of community "health" will be compared among the three sampled areas (0.1 m^2, 0.05 m^2 and 0.00785 m^2) and the two sieve sizes. The measures will include traditional metrics such as abundance, species richness, and relative abundance of potential indicator taxa, as well as multi-metric and multivariate indices of benthic condition that are being developed under Task 2 and Task 3. The outcomes will be assessed relative to sediment characteristics such as grain size, contamination level, and sample processing time.

4.3. Results. This task will produce a manuscript for publication that describes the advantages, disadvantages, and effects on ability to discriminate between polluted and unpolluted sites of each of the methods presently used in California. It will include recommendations on methods to be used for sampling benthos for assessing whether an area meets California's sediment quality objectives.

Task 5: Develop Taxonomy QA Procedures. The goal of taxonomy for applying benthic community SQOs is to identify all organisms to species because species identifications are necessary for calculating many of the benthic assessment measures presently in use. Several characteristics of benthic communities make accurate species identification and counts difficult. First, benthic communities are very diverse, including a broad range of animal phyla, classes, genera, and species. They can also be very abundant. It is not uncommon for infaunal samples from the nearshore sediments of California to contain over 200 species and several thousand individuals. The state of taxonomic science further compounds the challenge. The taxonomy of many groups is imperfectly known, and few taxonomists have the knowledge and experience necessary to accurately identify infauna from California's nearshore waters.

Repeatability of identifications and counts is an important consideration when more than one person identifies individuals of the same group of organisms. It is also an important consideration when sampling is repeated over time because unrecognized name changes can increase the apparent amount of change over time.

These challenges can be met and the data used for evaluating benthic community condition can be made more consistent by implementing a well-documented quality assurance (QA) program. Successful QA programs for standardizing identifications by multiple taxonomy teams have been developed in California (SCBPP, Bight'98, WEMAP) and will be used as the starting point for developing a statewide QA program for benthic community assessments of bays and estuaries.

5.1. Develop QA program. A taxonomy QA program will be developed that addresses three key topics: (1) sample sorting, (2) identifications and counts, and (3) taxonomist qualification. The QA program will be incorporated into a methods manual for benthic community assessment that will be produced as part of another element of the SQO project (see project overview). The characteristics of the QA program for each of the key topics are described below.

Sample sorting. Sorting is the process by which the organisms in a benthic sample are removed from the organic and inorganic residues that compose the sediments. The organisms are sorted into broad taxonomic categories for subsequent taxonomic analysis. Sorting is an exacting and time consuming task that requires close attention for prolonged periods and it is not unusual for this task to take several times longer than all other steps in the process. Despite its demands, it is essential that sorting be accurate. Incomplete sorting will result in unrecoverable undercounts of many organisms and species.

QA measures for sample sorting will include standard operating procedures (SOP) detailing and regulating the steps in the process, explicit quality objectives (e.g., removal of >95% of all organisms), quality control (QC) procedures that demonstrate compliance with the objectives, and remediation procedures in the event of failure to achieve quality objectives. A program of quality assurance for infaunal sample sorting with these elements will be developed for the sediment quality objectives program.

Identification and Counts. The goal of taxonomic analysis is accurate species-level identification and counts of all organisms collected. Each species must be recognized on each occasion it is encountered. If multiple taxonomists are involved in sample analysis (a typical situation), all taxonomists must treat each species similarly and use consistent nomenclature. Inconsistencies that affect data quality take several forms: miscounts, misidentifications, overlooked specimens, and differences in taxonomic level of identification and/or and nomenclature (Ranasinghe *et al.* 2003, Stribling *et al.* 2003).

QA measures to assure the accuracy and repeatability of identifications and counts will include sample processing SOPs, nomenclatural standards, voucher specimen collections and reviews, explicit quality objectives for counts, number of species, and identification accuracy, and QC procedures designed to assess compliance with quality objectives and characterize data quality. In addition, because taxonomic analysis is largely non-destructive, the processed samples themselves have considerable value as a QA resource and should be conserved whenever possible. A program of quality assurance for taxonomic analysis addressing these elements will be developed for the sediment quality objectives program. The program will address the issue of taxonomic sufficiency, identifying the levels of taxonomic identification necessary for different taxonomic groups and types of organisms to be eliminated prior to data analysis.

Schedule

Task	Activity or Deliverable	Completion Date
1: Identify Naturally Occurring Assemblages	1.1. Assemble database 1.2. Analyze Data 1.3. Preliminary Report	September 2004 December 2004 January 2005
2: Refine Existing Benthic Tools	2.1 Assemble Database 2.2 Analyze Data 2.3 Results	November 2004 January 2005 February 2005
3: Compare and Evaluate Benthic Tools	3.1 Assemble Database (Tasks 2 & 3 share a common database) 3.2 Analyze Data 3.3 Preliminary Report on Benthic Indicator Application Strategy	November 2004 May 2005 June 2005
4: Evaluate Field Sampling Methods	4.1 Collect Data and Assemble Database 4.2 Data Analysis 4.3 Results	February 2005 May 2005\ June 2005
5: Develop Taxonomy QA Procedures	5.1 Develop QA program	June 2005

Literature Cited

Alden, R.W., III, D.M. Dauer, J.A. Ranasinghe, L.C. Scott and R.J. Llanso. 2002. Statistical verification of the Chesapeake Bay benthic index of biotic integrity. *Environmetrics* **13**:473-498.

Anderson, B.S., J.W. Hunt, B.M. Phillips and R. Fairey. 2001. Sediment quality in Los Angeles Harbor, USA: a triad assessment. *Environ. Toxicol. Chem.* **20**:359-370.

Anderson, B.S., J.W. Hunt, B.M. Phillips, J. Newman, R.S. Tjeerdema, C.J. Wilson, G. Kapahi, R.A. Sapudar, M. Stephenson, H.M. Puckett, R. Fairey, J.M. Oakden, M. Lyons and S. Birosik. 1998. Sediment chemistry, toxicity, and benthic community conditions in selected water bodies of the Los Angeles region. California State Water Resources Control Board. Sacramento, CA.

Anderson, B.S., J.W. Hunt, S. Tudor, J. Newman, R.S. Tjeerdema, R. Fairey, J.M. Oakden, C. Bretz, C.J. Wilson, F. LaCaro, G. Kapahi, M. Stephenson, H.M. Puckett, J.W. Anderson, E.R. Long, T. Fleming and J.K. Summers. 1997. Chemistry, toxicity and benthic community conditions in sediments of selected southern California bays and estuaries. California State Water Resources Control Board. Sacramento, CA.

Bergen, M., D.B. Cadien, A. Dalkey, D.E. Montagne, R.W. Smith, J.K. Stull, R.G. Velarde and S.B. Weisberg. 2000. Assessment of benthic infaunal condition on the mainland shelf of Southern California. *Environ. Mon. and Assess.* **64**:421-434.

Bergen, M., S.B. Weisberg, D.B. Cadien, A. Dalkey, D.E. Montagne, R.W. Smith, J.K. Stull and R.G. Velarde. 1998. Southern California Bight 1994 Pilot Project Volume IV: Benthic Infauna. Southern California Coastal Water Research Project. Westminster, CA.

Bergen, M., S.B. Weisberg, R.W. Smith, D.B. Cadien, A. Dalkey, D.E. Montagne, J.K. Stull, R.G. Velarde and J.A. Ranasinghe. 2001. Relationship between depth, sediment, latitude, and the structure of benthic infaunal assemblages on the mainland shelf of southern California. *Mar. Biol.* **138**:637-647.

Bray, J.R. and J.T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs 27:325-349. *Ecol. Monogr.* **27**:325-349.

California State Water Resources Control Board. 1999. Consolidated Toxic Hot Spots Cleanup Plan: Draft Functional Equivalent Document. California State Water Resources Control Board. Sacramento, CA. New Series No. 6, Division of Water Quality.

Carlton, J.T. 1972. Current status of marine pollution research in the San Francisco Bay area. *A talk presented before the San Francisco Bay and Estuarine Association, April 20, 1972.*

Clifford, H.T. and W. Stephenson 1975. An introduction to numerical classification. New York, NY, Academic Press.

Dauer, D.M., J.A. Ranasinghe and S.B. Weisberg. 2000. Relationships between benthic community condition, water quality, sediment quality, nutrient loads, and land use patterns in Chesapeake Bay. *Estuaries* **23**:80-96.

Engle, V.D. and J.K. Summers. 1999. Refinement, validation, and application of a benthic condition index for Northern Gulf of Mexico estuaries. *Estuaries* **22**:624-635.

Engle, V.D., J.K. Summers and G.R. Gaston. 1994. A benthic index of environmental condition of Gulf of Mexico estuaries. *Estuaries* **17**:372-384.

Exponent. 2002. Technical memorandum 5: Phase 1 benthic macroinvertebrate data for the NASSCO and Southwest Marine detailed sediment investigation. Prepared for NASSCO and Southwest Marine. Bellevue, WA.

Fairey, R., C. Bretz, S. Lamerdin, J.W. Hunt, B.S. Anderson, S. Tudor, C.J. Wilson, F. LaCaro, M. Stephenson, H.M. Puckett and E.R. Long. 1996. Chemistry, toxicity and benthic community conditions in sediments of the San Diego Bay region. California State Water Resources Control Board. Sacramento, CA.

Fairey, R., J. Downing, C.A. Roberts, E. Landrau, J.W. Hunt, B.S. Anderson, C.J. Wilson, G. Kapahi, F. LaCaro, P. Michael, M. Stephenson and H.M. Puckett. 1998. Chemistry, toxicity, and benthic community conditions in sediments of the San Diego Bay region.

Final Addendum Report. California State Water Resources Control Board. Sacramento, CA.

Gibson, G.R., M.L. Bowman, J. Gerritsen and B.D. Snyder. 2000. Estuarine and coastal marine waters: bioassessment and biocriteria technical guidance. U. S. Environmental Protection Agency, Office of Water. Washington, DC. EPA 822-B-00-024.

Hale, S.S., J.F. Paul and J.F. Heltshe. 2004. Watershed landscape indicators of estuarine benthic condition. *Estuaries* **27**:283-295.

Hunt, J.W., B.S. Anderson, B.M. Phillips, R.S. Tjeerdema, K.M. Taberski, C.J. Wilson, H.M. Puckett, M. Stephenson, R. Fairey and J.M. Oakden. 2001. A large-scale categorization of sites in San Francisco Bay, USA, based on the sediment quality triad, toxicity identification evaluations, and gradient studies. *Environ. Toxicol. Chem.* **20**:1252-1265.

Hyland, J.L., W.L. Balthis, V.D. Engle, E.R. Long, J.F. Paul, J.K. Summers and R.F. Van Dolah. 2003. Incidence of stress in benthic communities along the US Atlantic and Gulf of Mexico coasts within different ranges of sediment contamination from chemical mixtures. *Environ. Mon. and Assess.* **81**:149-161.

Hyland, J.L., R.F. Van Dolah and T.R. Snoots. 1999. Predicting stress in benthic communities of southeastern U.S. estuaries in relation to chemical contamination of sediments. *Environ. Toxicol. Chem.* **18**:2557-2564.

Jacobi, M., R. Fairey, C.A. Roberts and E. Landrau. 1998. Chemical and biological measures of sediment quality and tissue bioaccumulation in the north coast region. California State Water Resources Control Board. Sacramento, CA.

Lance, G.H. and W.T. Williams. 1967. A general theory of classificatory sorting strategies. I. Hierarchical systems. *Computer Journal* **9**:373-380.

Llanso, R.J., L.C. Scott, J.L. Hyland, D.M. Dauer, D.E. Russell and F.W. Kutz. 2002. An estuarine index of biotic integrity for the mid-Atlantic region of the United States. II. Index Development. *Estuaries* **25**:1231-1242.

Long, E.R., C.B. Hong and C.G. Severn. 2001. Relationships between acute sediment toxicity in laboratory tests and abundance and diversity of benthic infauna in marine sediments: a review. *Environ. Toxicol. Chem.* **20**:46-60.

Long, E.R. and D.D. MacDonald. 1998. Recommended uses of empirically derived, sediment quality guidelines for marine and estuarine ecosystems. *Human and Ecological Risk Assessment* **4**:1019-1039.

Paul, J.F., K.J. Scott, D.E. Campbell, J.H. Gentile, C.S. Strobel, R.M. Valente, S.B. Weisberg, A.F. Holland and J.A. Ranasinghe. 2001. Developing and applying a benthic index of estuarine condition for the Virginian Biogeographic Province. *Ecological Indicators* **1**:83-99.

Ranasinghe, J.A., D.E. Montagne, R.W. Smith, T.K. Mikel, S.B. Weisberg, D.B. Cadien, R.G. Velarde and A. Dalkey. 2003. Southern California Bight 1998 Regional Monitoring Program: VII. Benthic Macrofauna. Southern California Coastal Water Research Project. Westminster, CA.

Ranasinghe, J.A., D.E. Montagne, S.B. Weisberg, M. Bergen and R.G. Velarde. 2003. Variability in the identification and enumeration of marine benthic invertebrate samples and its effect on benthic assessment measures. *Environ. Mon. and Assess.* **81**:199-206.

Rouse, G. 1994. Appendix 1. Report on the accuracy of identification of material in the BLM projects. Annual report to Minerals management Service and the national Biological Survey, Contract 14-35-001-30519: Disposition of Biological Specimens of the Outer Continental Shelf.

Smith, R.W., M. Bergen, S.B. Weisberg, D.B. Cadien, A. Dalkey, D.E. Montagne, J.K. Stull and R.G. Velarde. 2001. Benthic response index for assessing infaunal communities on the southern California mainland shelf. *Ecological Applications* **11**:1073-1087.

Smith, R.W., J.A. Ranasinghe, S.B. Weisberg, D.E. Montagne, D.B. Cadien, T.K. Mikel, R.G. Velarde and A. Dalkey. 2003. Extending the Southern California Benthic Response Index to Assess Benthic Condition in Bays. Southern California Coastal Water Research Program. Westminster, CA.

Southern California Coastal Water Research Project and Space and Naval Warfare Systems Center San Diego. 2004. Sediment assessment study for the mouths of Chollas and Paleta Creek, San Diego. Phase I Draft Report. San Diego Regional Water Quality Control Board, Commander Navy Region Southwest, City of San Diego. San Diego, CA.

Stribling, J.B., S.R. Moulton, II and G.T. Lester. 2003. Determining the quality of taxonomic data. *J. N. Am. Benthol. Soc.* **22**:621-631.

Thompson, B. and S. Lowe. 2004. Assessment of macrobenthos response to sediment contamination in the San Francisco Estuary, California, USA. *Environ. Toxicol. Chem.* **23:**2178-2187.

U.S. Environmental Protection Agency. 2004. Draft National Coastal Condition Report II. U.S. Environmental Protection Agency, Office of Research and Development. Washington, DC. EPA-620/R-03/002.

Van Dolah, R.F., J.L. Hyland, A.F. Holland, J.S. Rosen and T.R. Snoots. 1999. A benthic index of biological integrity for assessing habitat quality in estuaries of the southeastern USA. *Mar. Environ. Res.* **48**:269-283.

Weisberg, S.B., J.A. Ranasinghe, D.M. Dauer, L.C. Schaffner, R.J. Diaz and J.B. Frithsen. 1997. An estuarine benthic index of biotic integrity (B-IBI) for Chesapeake Bay. *Estuaries* **20**:149-158.